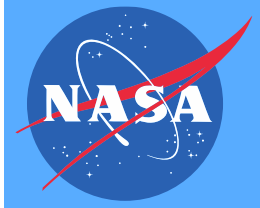


National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Tools and services for utilizing oceanographic data

Ed Armstrong, Thomas Huang, Eric Tauer
Physical Oceanographic DAAC
NASA Jet Propulsion Laboratory, Pasadena, CA

2017 CalCOFI Conference, San Diego, CA
Dec, 2017

PO.DAAC Organization

PO.DAAC Charter

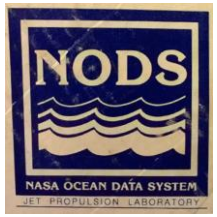
"To maintain NASA's oceanographic and climate data for future generations and to make these, and synergistic datasets, easily accessible, usable, and understandable for a broad and diverse set of data consumers."

* **Fundamental duty:** Maintain the archive (but… *active* archive; "easily accessible").
* **Fundamental value:** Make the data usable and understandable. *Facilitate science*!

PO.DAAC Legacy

PO.DAAC today

PO.DAAC (Physical Oceanography Distributed Active Archive Center)
1991 Joined DAAC system



***NODS** (NASA Ocean Data System)*
1987 Operational System

***PODS** (Pilot Ocean Data System)*
Pre-1986 pilot program

- *Distributing data from 10 NASA Projects and Missions*
- *Covering 6 key Physical Oceanography Parameters*
- *Project Holdings: 420* TB / 40,000,000 files*
- *Annual distribution of 500+ TB to over 46,000 unique users*
- *Public engagement with over 60,000 people annually*

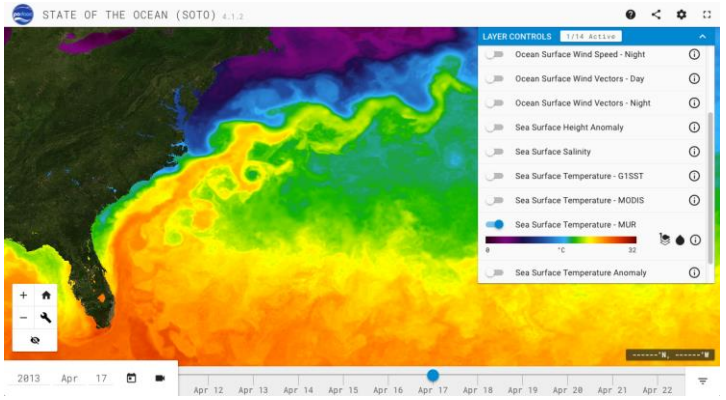
PO.DAAC Functional Areas

Data Management & Stewardship

Preserve NASA's data for the benefit of future generations

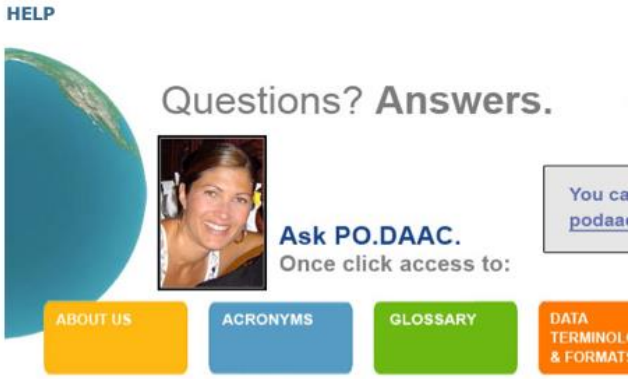
Data Access

Provide intuitive services to discover, select, extract and utilize data



Science Information Services

Provide a knowledgebase to help a broad user community understand and interpret satellite ocean data and related information

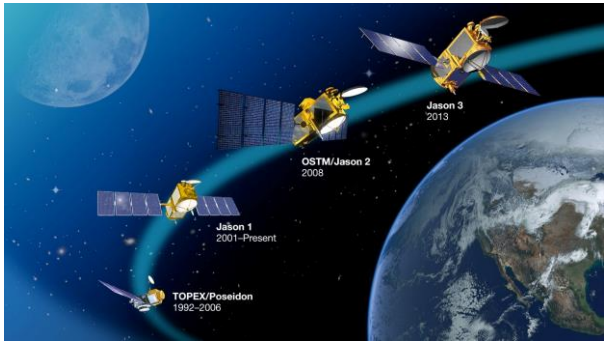


PO.DAAC Datasets

NASA Missions & Projects

Seasat, TOPEX/Poseidon, Jason-1, NSCAT, SeaWinds on ADEOS-II, QuikSCAT, GRACE, GHRSST, SPURS, MEaSUREs, Aquarius, CYGNSS, GRACE-FO (2017)

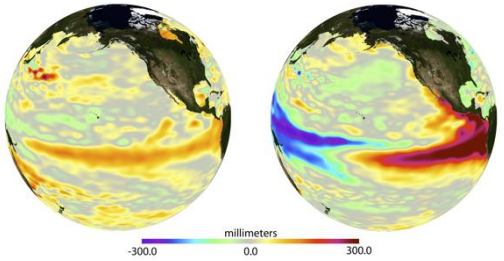
Upcoming: COWVR, *AirSWOT, SWOT, GRACE-2*

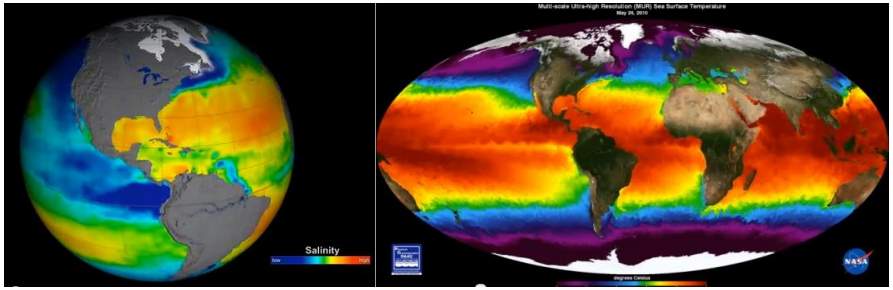


Ocean & Climate Community Driven

Value-added datasets in support of NASA programs

Gravity
Ocean Circulation & Currents
Ocean Surface Salinity
Ocean Surface Topography
Ocean Vector Winds
Sea Surface Temperature
Hydrology
Ocean Color
Sea Ice





PO.DAAC Partnerships

EUMETSAT

Royal Netherlands Meteorological Institute
Ministry of Transport, Public Works and Water Management

Earth & Space Research
A Nonprofit Research Institute

THE STATE UNIVERSITY OF NEW JERSEY RUTGERS

OSU Oregon State University

Center for Ocean-Atmospheric Prediction Studies

NASA

AQUARIUM OF THE PACIFIC

Ifremer

COAPS

FLORIDA STATE UNIVERSITY 1851

MICHIGAN

Massachusetts Institute of Technology

NOAA
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
U.S. DEPARTMENT OF COMMERCE

miami

WOODS HOLE OCEANOGRAPHIC INSTITUTION 1930

The EUMETSAT Network of Satellite Application Facilities

OSI SAF
Ocean and Sea Ice

USF
UNIVERSITY OF SOUTH FLORIDA

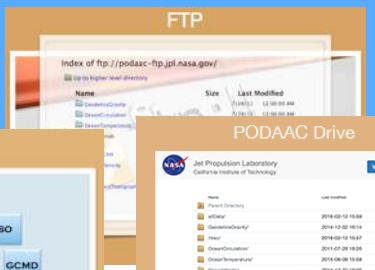
Scripps

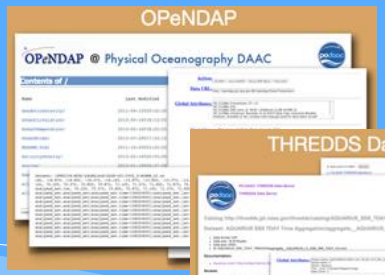
PO.DAAC Mission Portfolio

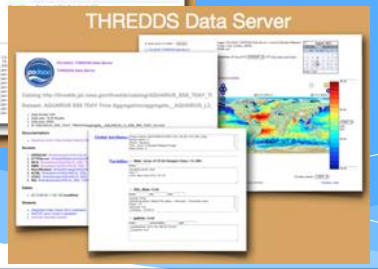
Parameter	Retired/Retiring	Ongoing	New	Future
SSS	Aquarius	SPURS I	SPURS II SMAP-SSS	
SST	GHRSST (GDS1)	GHRSST (GDS2) MODIS GOES VIIRS	Sentinel-3a	Sentinel-3b VIIRS/JPSS-1
Gravity	GEOS-3	GRACE		GRACE-FO
OST	TOPEX/Poseidon Jason-1	Jason-2/OSTM	Jason-3 Sentinel-3a OMG	SWOT Sentinel-3b Sentinel-6/Jason-CS
OVW	SeaSat NSCAT AMSR-E Seawinds on ADEOS-2	WindSat QuikSCAT RapidScat ASCAT	CYGNSS	MetOp-C COWVR
Ocean Circulation		OSCAR		

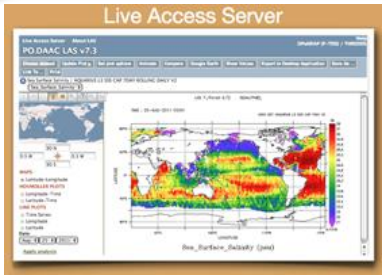
PO.DAAC Web, Tools and Services

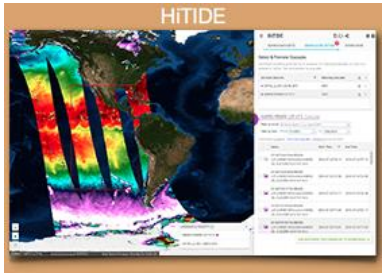
Protocols



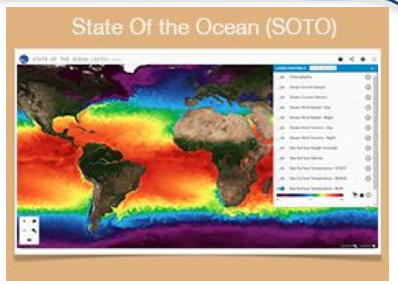


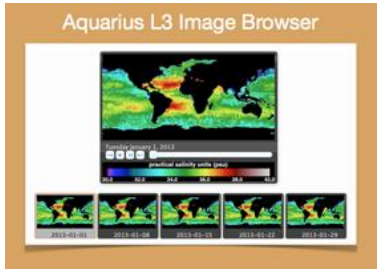




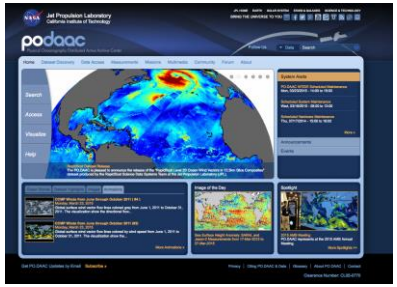


Subsetting



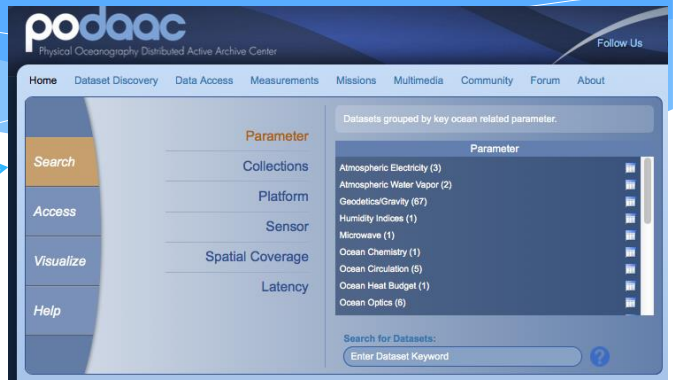


Visualization



Web Portal (Dataset Discovery)

PO.DAAC Web Portal



- A gateway to **Dataset Discovery** and **Data Access**
- Brings information together from diverse sources in a uniform way.
- Dynamic generations of content

- **Forums**
  - General Questions/FAQs
  - Data Recipes/Tutorials
  - Science Collaborative Forums

PO.DAAC Web Portal

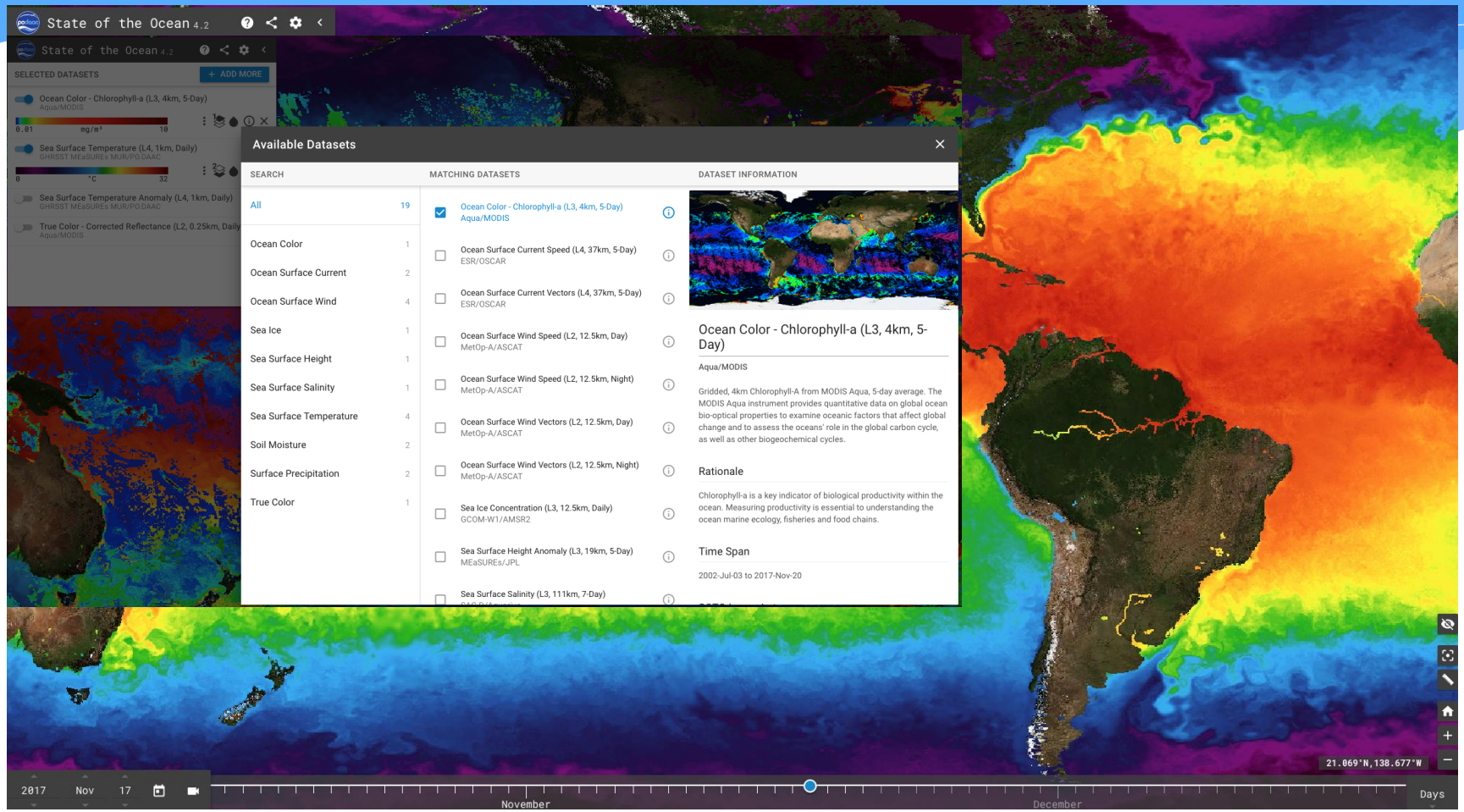
- **Dataset Discovery**
  - Faceted Browsing
  - Multi-level filtering
  - Keyword search
  - Dataset Information Page/DOI Landing Pages
  - Granule browsing through date tree

Dataset Information Page

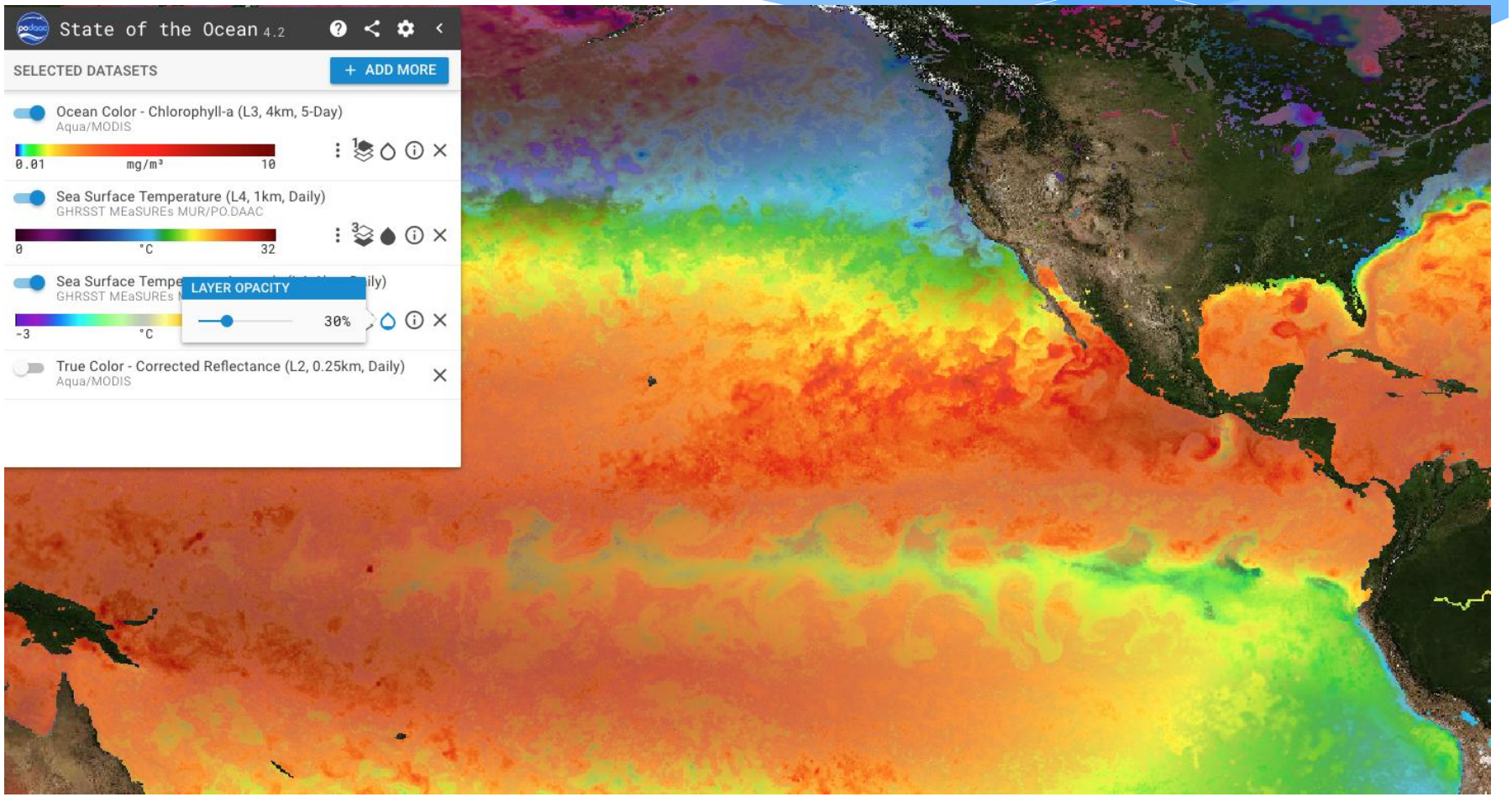
- **Information**
  - Dataset Metadata
- **Data Access**
  - Direct Access
  - Tools and Services
  - Read Software
- **Documentation**
  - Known Issues
- **Granule (File) Listing**
- **Citation**

SOTO Introduction

* State of the Ocean (SOTO) is a web client for multi layer visualization for ocean data of
  * Temperature
  * Wind
  * Sea Level
  * Salinity
  * Currents
  * Biology and others
* What else can SOTO be used for ?
  * Image overlay and animation
  * Color manipulation
  * Data interrogation
* Leverages tiled images created by the NASA Global Imagery Browse System (GIBS)


State of the Ocean 4.2
SELECTED DATASETS
+ ADD MORE
Ocean Color - Chlorophyll-a (L3, 4km, 5-Day)
Aqua/MODIS
0.01 mg/m³ 10
Sea Surface Temperature (L4, 1km, Daily)
GHRSST MEaSUREs MUR/PO.DAAC
0 °C 32
Sea Surface Temperature Anomaly (L4, 1km, Daily)
GHRSST MEaSUREs MUR/PO.DAAC
True Color - Corrected Reflectance (L2, 0.25km, Daily
Aqua/MODIS
Available Datasets
SEARCH
MATCHING DATASETS
DATASET INFORMATION
All 19
Ocean Color 1
Ocean Surface Current 2
Ocean Surface Wind 4
Sea Ice 1
Sea Surface Height 1
Sea Surface Salinity 1
Sea Surface Temperature 4
Soil Moisture 2
Surface Precipitation 2
True Color 1
Ocean Color - Chlorophyll-a (L3, 4km, 5-Day)
Aqua/MODIS
Ocean Surface Current Speed (L4, 37km, 5-Day)
ESR/OSCAR
Ocean Surface Current Vectors (L4, 37km, 5-Day)
ESR/OSCAR
Ocean Surface Wind Speed (L2, 12.5km, Day)
MetOp-A/ASCAT
Ocean Surface Wind Speed (L2, 12.5km, Night)
MetOp-A/ASCAT
Ocean Surface Wind Vectors (L2, 12.5km, Day)
MetOp-A/ASCAT
Ocean Surface Wind Vectors (L2, 12.5km, Night)
MetOp-A/ASCAT
Sea Ice Concentration (L3, 12.5km, Daily)
GCOM-W1/AMSR2
Sea Surface Height Anomaly (L3, 19km, 5-Day)
MEaSUREs/JPL
Sea Surface Salinity (L3, 111km, 7-Day)
Ocean Color - Chlorophyll-a (L3, 4km, 5-Day)
Aqua/MODIS
Gridded, 4km Chlorophyll-A from MODIS Aqua, 5-day average. The MODIS Aqua instrument provides quantitative data on global ocean bio-optical properties to examine oceanic factors that affect global change and to assess the oceans' role in the global carbon cycle, as well as other biogeochemical cycles.
Rationale
Chlorophyll-a is a key indicator of biological productivity within the ocean. Measuring productivity is essential to understanding the ocean marine ecology, fisheries and food chains.
Time Span
2002-Jul-03 to 2017-Nov-20
21.069°N,138.677°W
2017 Nov 17
November
December
Days

SST Imagery Overlay



Use case: Gulf of Tehuantepec gap winds

* Recent Gulf of Tehuantepec wind event
* https://podaac-tools.jpl.nasa.gov/soto/#b=BlueMarble_ShadedRelief_Bathymetry&l=jpl_l4_mur_ssta___ssta___36000_x_18000___daynight(la=true),ascata_l2_coastal___wind_dir_color_day_(la=true),ascata_l2_coastal___wind_dir_color_night_(la=true),GHRSST_L4_MUR_Sea_Surface_Temperature,MODIS_Aqua_CorrectedReflectance_TrueColor,modis_aqua_l3_chla_daily_4km_l___chlorophyll_a___8640_x_4320___daynight&ve=-111.072721483934,7.67652268408321,-85.44968512850639,21.6070506552272 6&pl=false&pb=false&d=2017-11-15&ao=false&as=2017-10-25&ae=2017-11-08&(censored-word)=1/day&afr=500&tlr=days

HiTIDE (*High-level Tool for Interactive Data Extraction*) Introduction

* A Level-2 (swath) granule (file) subsetting tool allows a user to inspect, extract and download granules in a region of interest.

* **Key features:**

- *File selection*

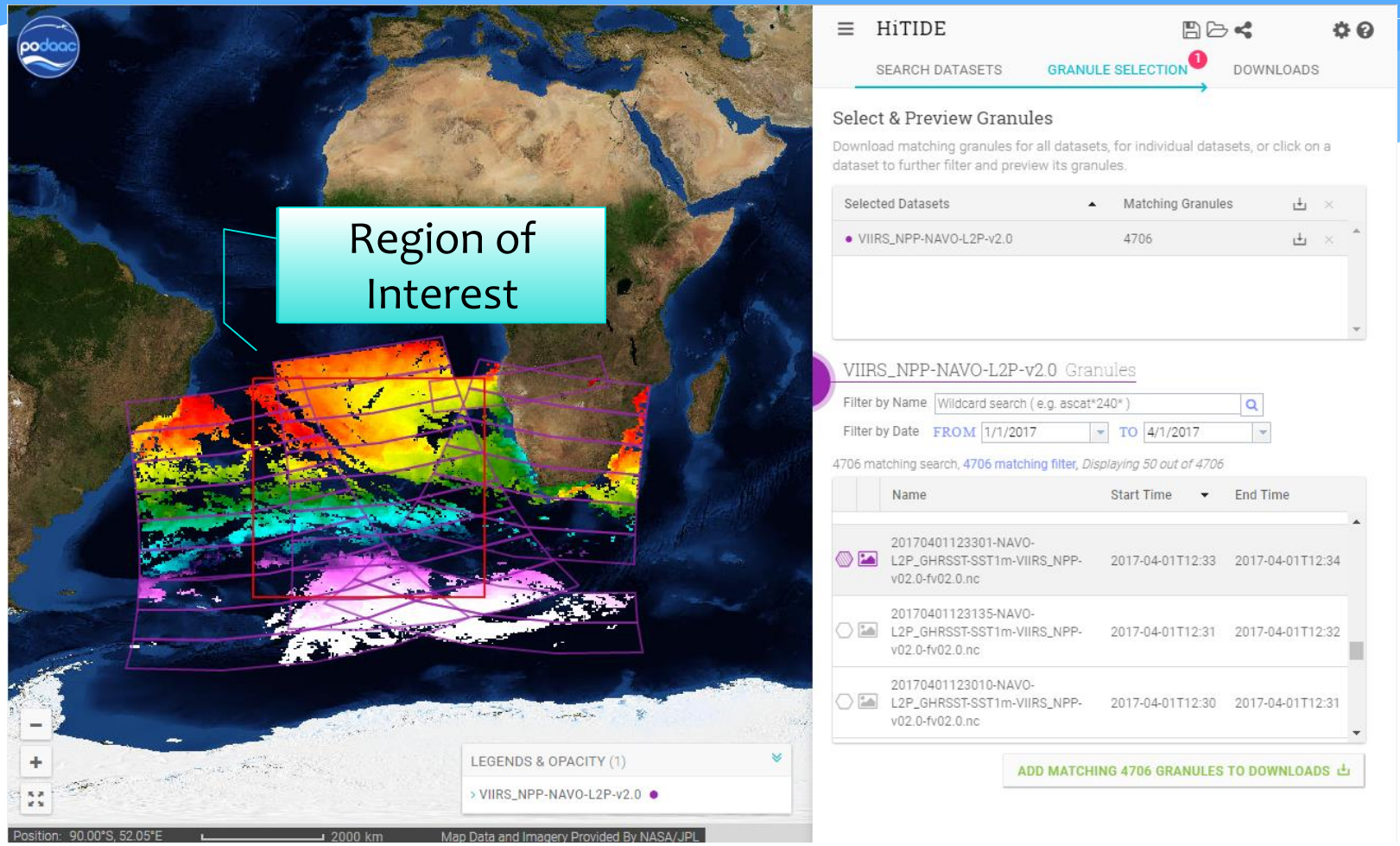
1. Filtering of datasets by Variable, Sensor, and Provider
2. Bounding by the time range and region of interest

- *Visual inspection of selected granules*

1. Display of orbit track of granules
2. Display of 2-D Level-2 (swath) granule image

- *Download: Files in the region of interest can be downloaded and aggregated in native format and saved for further scientific analysis.*

HiTIDE


Region of Interest
HiTIDE
SEARCH DATASETS
GRANULE SELECTION
DOWNLOADS
Select & Preview Granules
Download matching granules for all datasets, for individual datasets, or click on a dataset to further filter and preview its granules.
Selected Datasets
Matching Granules
VIIRS_NPP-NAVO-L2P-v2.0
4706
VIIRS_NPP-NAVO-L2P-v2.0 Granules
Filter by Name
Wildcard search (e.g. ascat*240*)
Filter by Date
FROM 1/1/2017
TO 4/1/2017
4706 matching search, 4706 matching filter, Displaying 50 out of 4706
Name
Start Time
End Time
20170401123301-NAVO-L2P_GHRSST-SST1m-VIIRS_NPP-v02.0-fv02.0.nc
2017-04-01T12:33
2017-04-01T12:34
20170401123135-NAVO-L2P_GHRSST-SST1m-VIIRS_NPP-v02.0-fv02.0.nc
2017-04-01T12:31
2017-04-01T12:32
20170401123010-NAVO-L2P_GHRSST-SST1m-VIIRS_NPP-v02.0-fv02.0.nc
2017-04-01T12:30
2017-04-01T12:31
ADD MATCHING 4706 GRANULES TO DOWNLOADS
LEGENDS & OPACITY (1)
VIIRS_NPP-NAVO-L2P-v2.0
Position: 90.00°S, 52.05°E
2000 km
Map Data and Imagery Provided By NASA/JPL

Web Services

* Web Services
* PODAAC also provides the efficient machine-to-machine communication and data transfer via script programming using its web services APIs including dataset discovery, granule discovery, granule imaging and subset, and metadata extraction. The table on the right showing the available Web Services at PO.DAAC. A typical PO.DAAC Web Service request is of the following URL template, which aligns with the standard “Open Search Protocol”:
* https://podaac.jpl.nasa.gov/ws/service?parameters

Web Services

The following is the list of available PO.DAAC Web Services

Name	Description
Dataset Metadata	Dataset metadata service retrieves the metadata of a dataset on PO.DAAC's dataset catalog using the following parameters: datasetId, shortName, and format.
Dataset Search	Dataset Search service searches PO.DAAC's dataset catalog, over Level 2, Level 3, and Level 4 datasets, using the following parameters: datasetId, shortName, startTime, endTime, bbox, and others.
Dataset Variables	Provides list of dataset variables.
Granule Metadata	Granule metadata service retrieves the metadata of a granule on PO.DAAC's catalog using the following parameters: format and other optional parameters.
Granule Search	Search Granule does granule searching on PO.DAAC level 2 swath datasets (individual orbits of a satellite), and level 3 & 4 gridded datasets (time averaged to span the globe). The following parameters are supported: datasetId, startTime, endTime, bbox, and others.
Granule Preview	The PODAAC preview Image service retrieves pre-generated preview images for selected granules. This service uses a template provided by the Granule Search service and, therefore, Granule Preview must be preceded by Granule Search.
Granule Subset	Subset Granule service allows users to Submit subset jobs. Use of this service should be preceded by a Granule Search in order to identify and generate a list of granules to be subsetted.
Subset Status	Subset Granule Status service allows users to check the status of submitted subset job.

Examples

This example python script shows an entire subsetting workflow using the PO.DAAC Web Services, exercising several of the services: datasets are searched (Dataset Search); variables in the found datasets are identified (Dataset Variables); granules meeting search criteria are found (Granule Search); and the identified granules are subsetted (Granule Subset). Before running this script, email_address variable need to be changed to a valid email address to receive the result. Use the following command to run the script

```
python cws_example.py
```

Emerging Missions-- SWOT

* NASA PO.DAAC is a data archive center that will be distributing Surface Water and Ocean Topography (SWOT) data after launch in 2021. **This data will be useful for anyone interested in the water cycle (oceans, lakes, reservoirs, rivers, ice, coasts, etc.).** Please take a survey (~10-15 min) to help us learn how you could use the data so that we can provide the appropriate tools and services.
*
* Tinyurl.com/swotsurvey2

Emerging Technologies

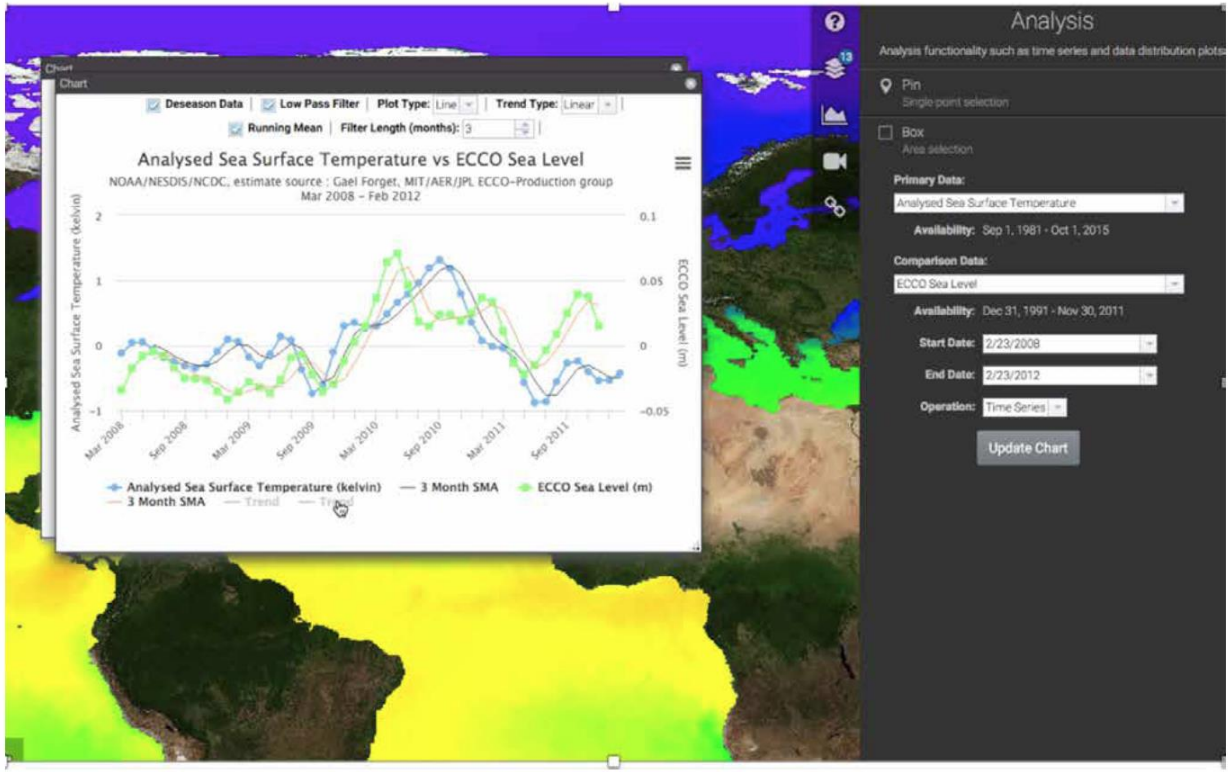
* **VQSS** *(Virtual Quality Screening Service)*
  * VQSS has developed an infrastructure to expose, apply, and extract quality screening information
* **MUDROD** *(Mining and Utilizing Dataset Relevancy from Oceanographic Dataset)*
  * MUDROD improvesweb-based search relevancy and ranking for NASA earth science products served by the PO.DAAC
* **DOMS** *(Distributed Oceanographic Matchup Service)*
  * Matchup service between Level 2 satellite data and in situ observations
* **OceanXtremes**
  * Analytic services for analysis and correlating long time series oceanographic satellite data including anomaly (outlier) detection

OceanWorks

National Aeronautics and Space Administration

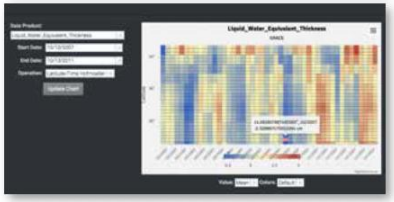
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Analyze Sea Level On-The-Fly *https://sealevel.nasa.gov*



Sea Level Change - Data Analysis Tool

- Visualizations
- Time Series
- Deseason
- Data Comparison
- Latitude/Time Hovmöller
- Etc.

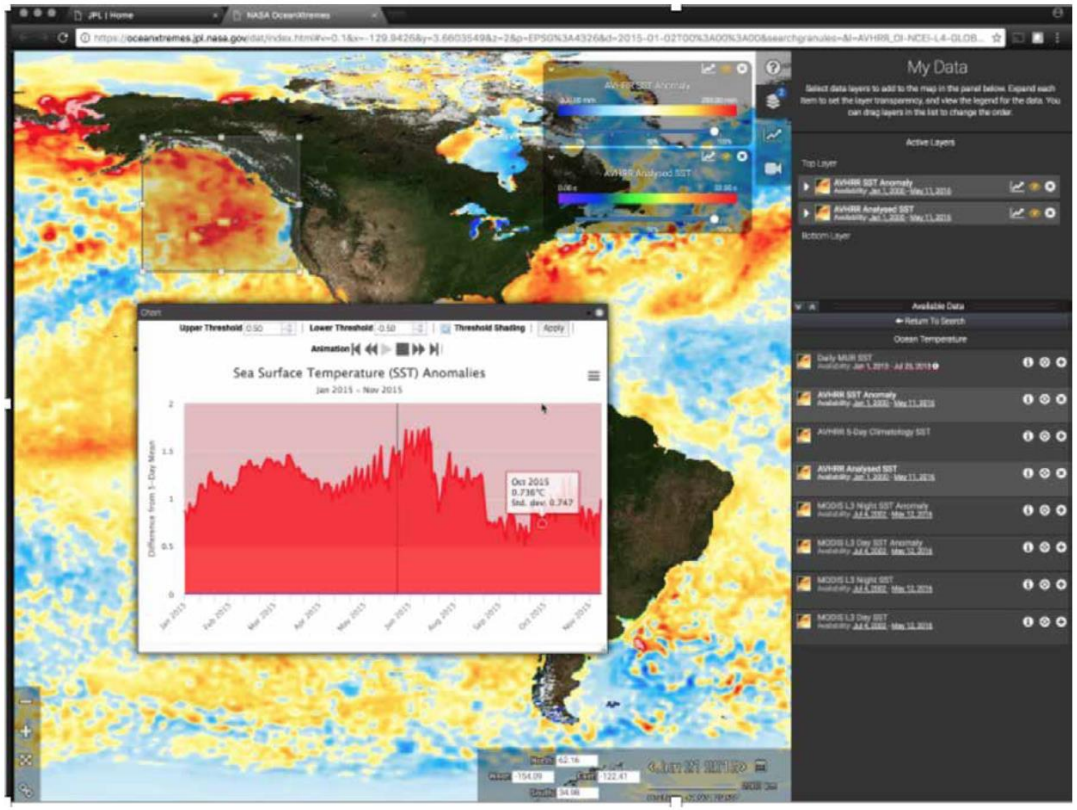


OceanWorks

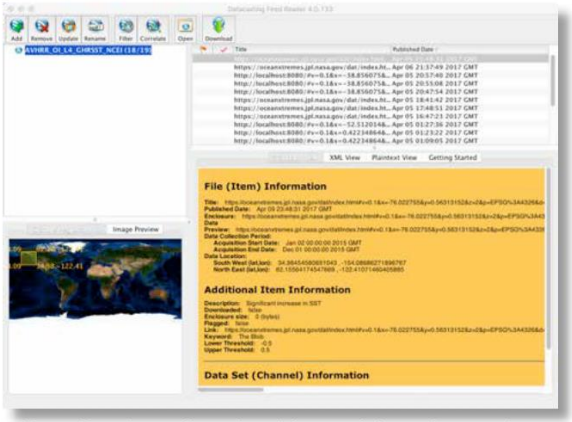
National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Analyze Ocean Anomaly



- **Visualize** parameter
- **Compute** daily differences against climatology
- **Analyze** time series area averaged differences
- **Replay** the anomaly and visualize with other measurements
- **Document** the anomaly
- **Publish** the anomaly

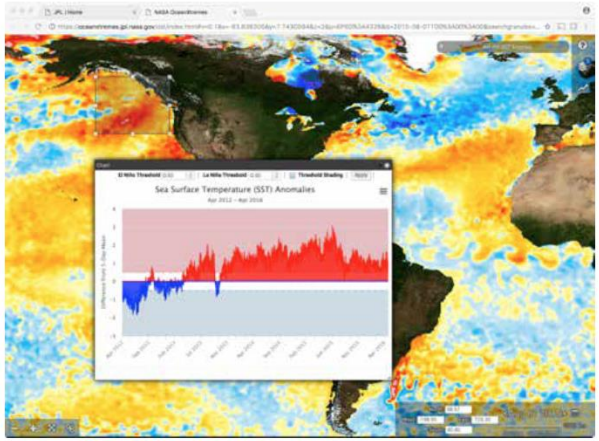


DataCasting client is able to pickup Anomaly Cast published by OceanXtremes

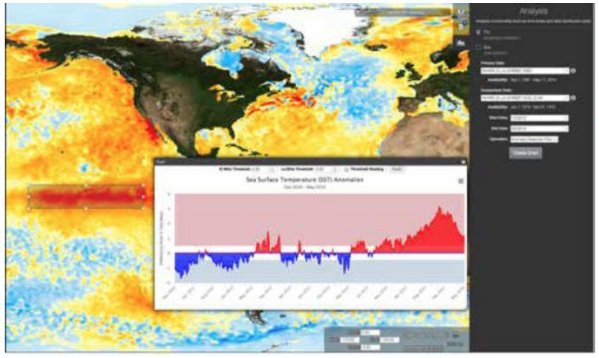
OceanWorks

National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

More Anomalies



"The Blob"



El Niño 3.4 regional signal

Recreated identification of "The Blob"

- **The Blob** is the name given to a large mass of relatively warm water in the Pacific ocean off the coast of North America. It was first detected in late 2013 and continued to spread throughout 2014 and 2015.
- SST anomaly = SST – SST Climatology at each location to compare with standard deviation - Chelle Gentemann, Senior Scientist at Earth & Space Research

Recreated the El Niño 3.4 regional signal

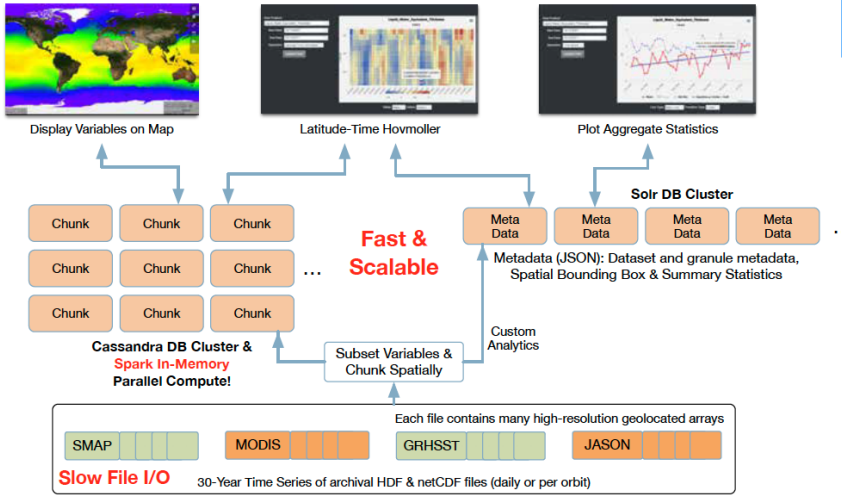
- **El Niño** is a phenomenon in the equatorial Pacific Ocean characterized by a five consecutive 3-month running mean of sea surface temperature (SST) anomalies in the Niño 3.4 region that is above (below) the threshold of +0.5°C (-0.5°C). This standard of measure is known as the Oceanic Niño Index (ONI).
- https://www.ncdc.noaa.gov/teleconnections/enso/indicators/sst.php

OceanWorks

National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Scalable Data Analytic Solution

- NEXUS is a data-intensive analysis solution using a new approach for handling science data to enable large-scale data analysis
- Streaming architecture for horizontal scale data ingestion
- Scales horizontally to handle massive amount of data in parallel
- Provides high-performance geospatial and indexed search solution
- Provides tiled data storage architecture to eliminate file I/O overhead
- A growing collection of science analysis webservices using Apache Spark: parallel compute, in-memory map-reduce framework
- Pre-Chunk and Summarize Key Variables
  - Easy statistics instantly (milliseconds)
  - Harder statistics on-demand using Spark (in seconds)
  - Visualize original data (layers) on a map quickly (Cassandra store)
- **Algorithms** – Time Series | Latitude/Time Hovmöller| Longitude/Time Hovmöller| Latitude/Longitude Time Average | Area Averaged Time Series | Time Averaged Map | Climatological Map | Correlation Map | Daily Difference Average



Two-Database Architecture

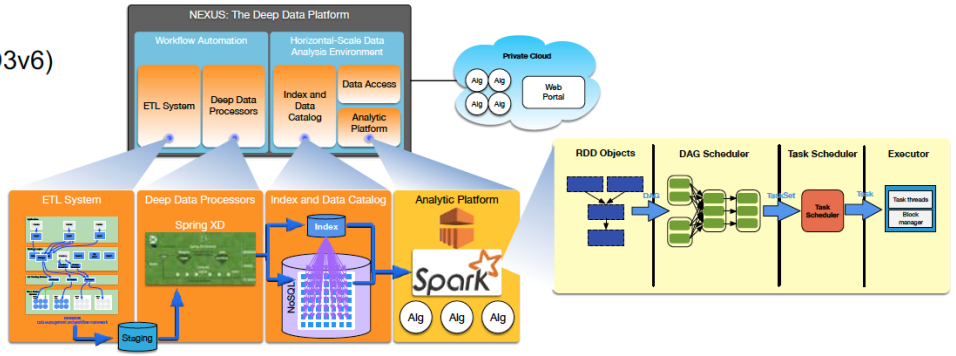
Open Source: Apache License 2
https://github.com/dataplumber/nexus

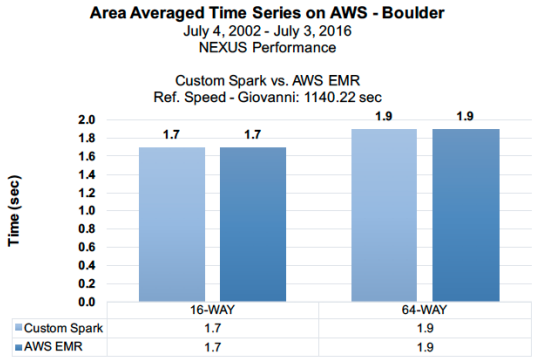
OceanWorks

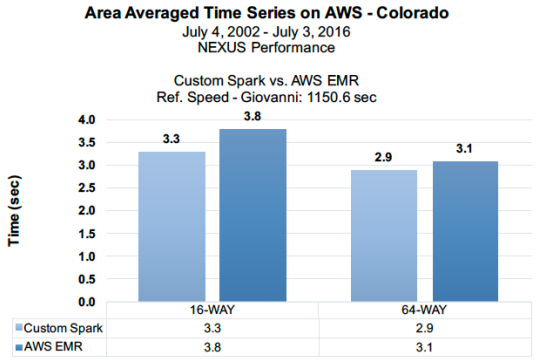
National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

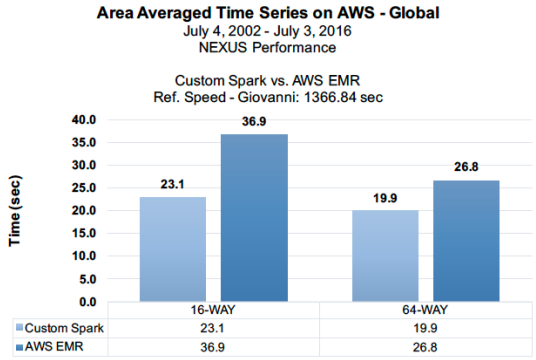
NEXUS Performance: Custom Spark vs. AWS EMR

Dataset: MODIS AQUA Daily
Name: Aerosol Optical Depth 550 nm (Dark Target) (MYD08_D3v6)
File Count: 5106
Volume: 2.6GB
Time Coverage: July 4, 2002 – July 3, 2016







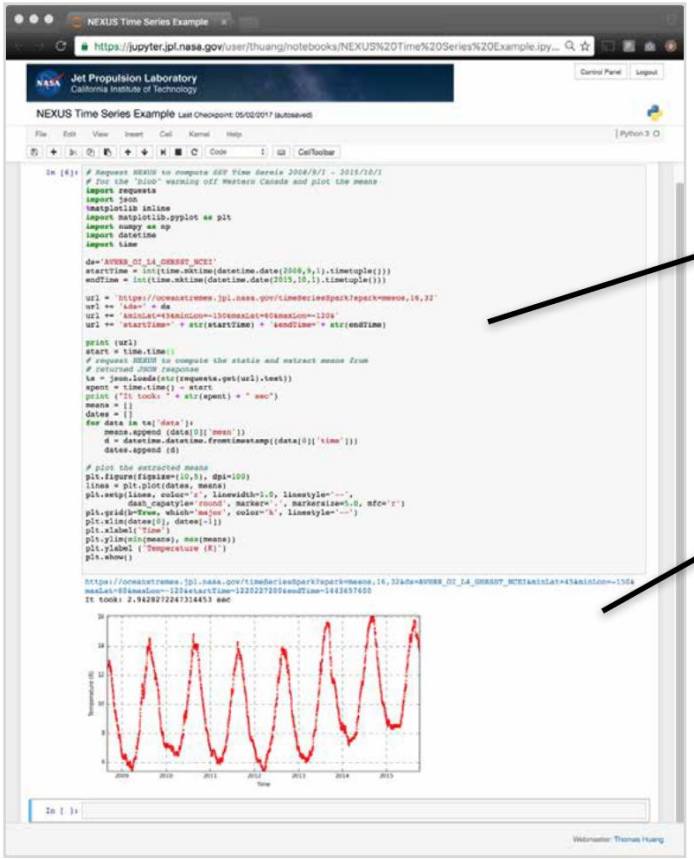


OceanWorks

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Enable Science without File Download

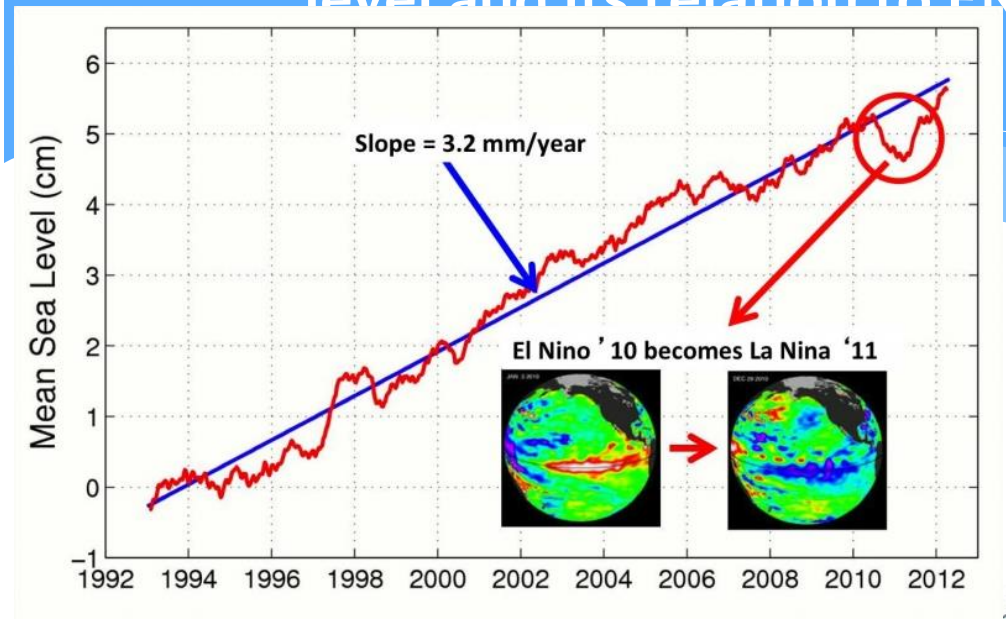



```
# Request NEXUS to compute SST Time Series 2008/9/1 - 2015/10/1
# for the "blob" warming off Western Canada and plot the means
…
ds='AVHRR_OI_L4_GHRSST_NCEI'

url = … # construct the webservice URL request


# make request to NEXUS using URL request
# save JSON response in local variable
ts = json.loads(str(requests.get(url).text))

# extract dates and means from the response
means = []
dates = []
for data in ts['data']:
    means.append (data[0]['mean'])
    d = datetime.datetime.fromtimestamp((data[0]['time']))
    dates.append (d)

# plot the result
…
```

```
https://oceanxtremes.jpl.nasa.gov/timeSeriesSpark?spark=mesos,16,32&ds=AVHRR_OI_L4_GHRSST_NCEI&minLat=45&minLon=-150&maxLat=60&maxLon=-120&startTime=1220227200&endTime=1443657600

It took: 2.9428272247314453 sec
```

backups

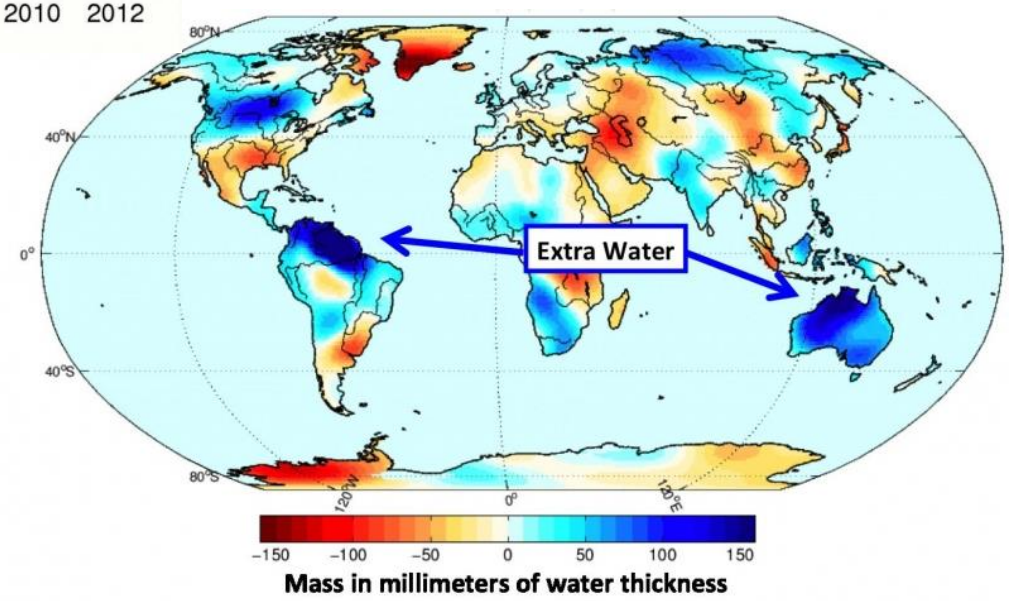
GRACE – Uncovering the 2010-11 decline in global mean sea level and its relation to ENSO (October, 2012)



In mid 2010 to mid 2011, global mean sea level dropped by ~5 mm

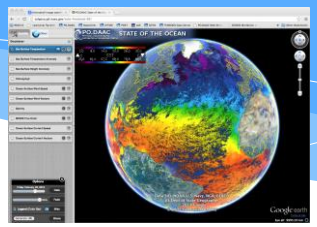
GRACE observations revealed that the sudden drop in sea level was related to a temporary loss of ocean water that was redistributed by evaporation, rain, and winds to the continents.

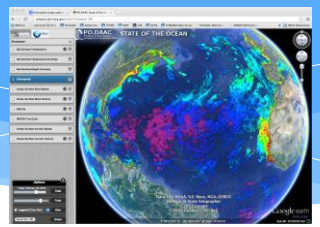
High precipitation events occurred during the 2010-11 La Niña and brought a massive amount of water to Australia and northern South America.



SOTO user story #1: hurricane impact on fishery

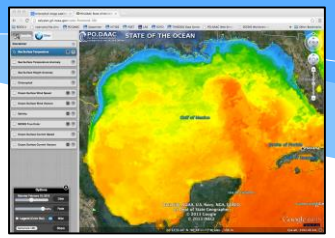


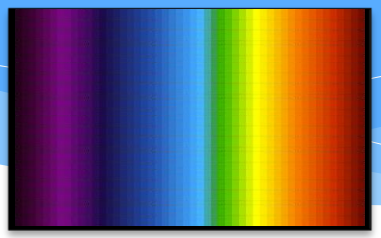




* A fishery company hears a tropical storm has been upgraded to a category 1 hurricane in the North Atlantic Ocean. The company needs to understand how it will impact their shipping fleet. Also, they want to quickly assess if there is an ocean response and if so, by what magnitude (e.g., significant regional decreases in SST, areas of relatively high chlorophyll-a concentrations) so they can approximate expected fish catchments.

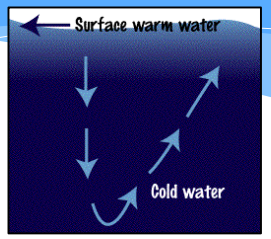
SOTO user story #2: Gulf of Mexico research

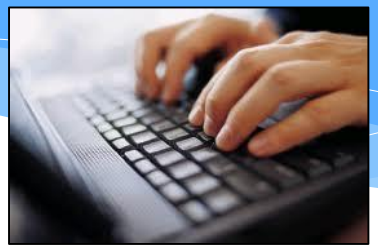




* A scientist wants information on the current ocean conditions in the hypoxic zone of the Gulf of Mexico in summer. The scientist first runs a visualization to see if anything is out of the ordinary, then downloads data files for that particular time. The scientist knows temperatures in the Gulf of Mexico often exceed 29°C in summer and needs a color bar appropriate for the region.

SOTO user story #3: connecting scientists to students





* An ocean algal bloom event is happening off the nearby coast. Analysis by a scientist using SOTO data sources such as chlorophyll indicates the event has been ongoing for 2 weeks and may be due to coastal upwelling. The scientist enters the algal bloom event into SOTO including a description of the analysis. A high school student using SOTO notices the algal bloom event, reads the description and submits a follow-up question which the scientist in turn replies to, generating a Q&A session.